DOCUMENT RESUME

ED 280 900 TM 870 263

AUTHOR Prell, JoAnn M.; Prell, Paul A.
TITLE Improving Test Scores--Teaching Test-Wiseness. A Review of the Literature. Research Bulletin.

INSTITUTION Phi Delta Kappa, Bloomington, IN. Center on Evaluation, Development, and Research.

PUB DATE Nov 86
NOTE 5p.
PUB TYPE Information Analyses (070) -- Collected Works - Serials (022)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Age Differences; Elementary Education; Intelligence Quotient; Literature Reviews; Racial Differences; Sex Differences; Socioeconomic Influences; Standardized Tests; Teacher Made Tests; *Test Coaching; *Testing Problems; Test Validity; *Test Wiseness

ABSTRACT

The role of testing in American education is becoming increasingly important in grouping students for educational purposes. This article reviews practices adopted by educators or administrators to improve student test scores. Stopgap measures for improving standardized test scores include (1) switching to easier tests; (2) eliminating certain populations from the testing; (3) use of coaching techniques; and (4) the teaching of test-wiseness. Teaching students how to take tests has several advantages over other methods used to inflate test scores. It requires a minimal amount of time, is relatively inexpensive, and is considered one of the more ethical procedures. The Millman taxonomy defines test-wiseness and breaks it down into two major areas--elements either independent or dependent of the test constructor or purpose. The material discussed in the taxonomy applies to teacher-made-tests and to standardized tests. The principles of test-wiseness can be taught using any of a variety of means: supervised study and lecture, independent reading, programmed instruction, and instruction through the mail. Test-wiseness has also been studied in relation to age, sex, IQ, race, and socioeconomic level. The conclusion is made that one of the most important aspects of test-wiseness deals with its relationship to test validity. (LMO)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED280900

Research Bulletin

Center on Evaluation, Development and Research

Phi Delta Kappa
November 1986 5

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Kliewer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Improving Test Scores – Teaching Test-Wiseness

A Review of the Literature

In 1983, school systems purchased an estimated $500 million worth of commercial standardized achievement tests – an estimate that does not even include IQ, diagnostic, or most minimum competency tests.[1] Over 200 million achievement test forms alone were used.[2] The role of testing in American education is becoming increasingly important. Tests are being used to determine grouping of students for classroom instruction, placement of transfer students, identification of educationally maladjusted children, students' vocational or educational goals, assessment of student achievement, assessment of teacher effectiveness, evaluation of programs, and much more.

The testing movement has continued to grow over the past 60 years and has most recently added a new instrument, the competency test, which is being used to determine a student's readiness for advancement or promotion. Competency test results are used to determine which college a student can attend or if he can attend college at all. Professions such as law, medicine, and accounting have had competency tests – bar exams or their counterparts – for years. Now, many educators are considering the use of a similar exam. "The temptation to turn to tests to measure individual readiness or achievement is actually growing on the American educational scene. Students will undoubtedly continue to be required to take tests of all kinds in the future. It is unlikely that the number and/or importance of tests will be diminished."[3]

The growth of the testing movement has been accompanied by "a variety of related interests, judgmental commentaries, and practical reactions."[4] With ever-increasing emphasis on standardized testing in the schools we are seeing practices develop that involve ethical haziness. "Every school board member and superintendent wants his school to come out on top. Lately, though, earning laurels because of student test scores has become an increasingly precarious, elusive career."[5] When school officials suffer criticism over poor test performance, they may resort to questionable practices such as boosting scores.

Stopgap Measures

In response to the pressure on educators or administrators to improve the standing of their class, school, or district, several approaches to improving standardized test scores have emerged. These include switching to easier tests, eliminating certain populations from the testing, use of coaching techniques, and the teaching of test-wiseness.

David Savage, an education writer for the *Los Angeles Times* argues, "The quickest way to raise scores is to switch to an easier test."[6] "In 1981, when the Cleveland schools substituted a reputedly easier test for the test they had been using, the number of students with scores in the average to above-average range jumped 10 percent."[7] Somehow superintendents and other school administrators claim to be able to rank the tests from easiest to most difficult. Although the rankings are based on hearsay rather than substantiated fact, the rankings remain the basis for this questionable method of improving test results. The extent to which this method is used is unknown, but one "Ohio superintendent said that the substantial cost of reviewing and adopting new tests might keep this method in check."[8]

More common than switching tests is the practice of making sure either that certain students do not take a specific test or that the test scores of some students are not counted with the general test results. Charges of score inflation are not uncommon. This seems especially true in big cities, where school systems are under pressure to improve their students' standing compared to national norms. "Dr. William Corbett, principal of Lowell Elementary School in Watertown, Massachusetts, suggests that principals or superintendents can manipulate their scores significantly by allowing the exclusion, for example, of special education students from the testing. This method was also cited in a *New York Post* article describing New York City's 1980 methodology for reporting score information, which excluded bilingual, truant, and special education students."[9]

A third method for improving overall test scores is coaching for the test. While there is no universally accepted definition for "coaching," the popular use of the term means training children to answer specific types of questions and providing the information required by a specific test. In 1977, the schools in Washington, D.C. came in "dead last" in the race for standardized achievement. "James Guines, the Associate Superintendent for Instruction and the administrator in charge of the Washington school system's test preparation program, stated, 'We want all our students to do well. We'll even begin to tell them some of the things on the test.' "[10] Some educators call this teaching to the test. This practice is controversial and raises ethical concerns. Many testing directors draw

IM870263

BEST COPY AVAILABLE

a fine line between teaching the material to be covered on the test and teaching to the test.

Teaching Students How to Take Tests

The fourth method for improving test scores that has been studied recently also employs a form of "coaching." Whereas the previous method involved coaching with the content of the test, this method involves coaching students on how to take tests. Instruction in test taking skills, or test-wiseness as it is referred to by researchers, has several advantages over other methods used to inflate test scores. It requires a minimal amount of time, is relatively inexpensive, and is considered as one of the more ethical procedures. Evaluation Specialist Jarold Zacharius stated, "if you're going to be involved with these tests, it would be stupid not to engage in test-taking instruction. As far as I can tell, it's in no way an immoral thing to do."[11]

The topic of test-wiseness has a relatively short history in educational research. Thorndike wrote a brief descriptive treatise on the subject in 1951. Gibb provided the first thorough empirical investigation on test-wiseness in 1964. The work that is considered the "classic theoretical work in the area" was done by Millman, Bishop, and Ebel in 1965. The Millman definition and comprehensive taxonomy of test-wiseness have been cited in most research in this area.[12]

The Millman definition states that test-wiseness is "a subject's capacity to utilize the characteristics and formats of the test and/or the test-taking situation to receive a high score."[13] Millman also noted that "test-wiseness is logically independent of the examinee's knowledge of the subject matter for which the items are supposedly measures."[14]

The Millman taxonomy is divided into two major areas – elements either independent or dependent of the test constructor or purpose. Major strategies addressed under the heading of independent elements are 1) use of time, 2) avoidance of errors, 3) guessing, and 4) deductive reasoning. Dependent element strategies included consideration of intent and use of cues.[15]

Valerie Ford, in her work entitled *Everything You Wanted to Know About Test-Wiseness*, elaborates on the above taxonomy. Her new outline, "Selected Aspects of Test-Wiseness," is based on research findings since Millman's work. Strategies added in the new outline include instructions specific to item format, e.g., answering analogy questions and multiple-choice questions. She also discusses the advisability of answer changing.[16]

In analyzing the Millman taxonomy it should be noted that not all the items would be applicable to a particular testing situation. Also, it appears that the principles discussed are useful only with objective multiple choice tests. According to Sarnacki, however, "test-wiseness is not limited to objective tests. Certain aspects of the time-using and error avoidance strategies are applicable to tests more subjective in nature, such as essay examinations. Furthermore, knowledge of the test-constructor's intent and idiosyncrasies in test making and grading, will aid the examinee in maximizing scores regardless of question format. Test-wiseness abilities generalize to a number of situations."[17]

It might appear that much of the material discussed in the Millman taxonomy applies only to teacher-made tests and that standardized tests are immune to test-wiseness. However, research indicates that this is not the case. "A number of researchers have shown that training in various test-wiseness skills will improve test scores on subsequent standardized testing. The tests employed in these studies included: Metropolitan Readiness Test (Oakland, 1972), California Test of Basic Skills (Gaines and Jongsma, 1974), Metropolitan Achievement Test (Gross, 1976), Stanford Reading Test (Callenbach, 1973), and the Iowa Test of Educational Development (Ormvig, 1971)."[18] The preceding list of tests was prepared by Sarnacki. Jongsma and Warshauer, in their reviews of research, found even more standardized tests being used in the studies of test-wiseness. In particular the use of the Lorge Thorndike IQ Test and Otis Lennon IQ Test were noted.[19] Only one research project could be found that employed the use of the Iowa Test of Basic Skills. The research project was conducted by Diamond and Evans in 1972 and dealt with sixth graders.[20]

Whether the test is teacher-made or standardized, the particular subject matter being tested is not exempt from being influenced by test-wiseness. The extent to which this is evident is seen in the variety of subject areas involved in test-wiseness research. Sarnacki, in his review of research, has listed language and math, word knowledge and science, American history, political science, algebra, vocabulary, analogies, general information, and reading as subjects used for testing in test-wiseness research. Another study, conducted by Bajtelsmit in 1975, demonstrated more emphatically this independence of test-wiseness from content area. The testing instrument involved content-free (nonsensical) items, in which the correct answer could only be arrived at through the application of specific test-wiseness principles.[21]

The principles of test-wiseness can be taught using any of a variety of means. Research has shown success with supervised study and lecture, independent reading, programmed instruction, and instruction through the mail.[22]

Most researchers agree that students need not spend a lot of time learning test-wiseness in order to reap significant results. Jones and Ligon stated that most instructional units (lessons) used in research studies were no more than one or two hours in length.[23] In Samson's 1985 meta-analysis of 24 research studies, he found the largest mean effect size occurred when the length of treatment was from five to seven weeks.[24] Therefore, it would seem that instruction of from five to 14 hours would be most effective. Bangert-Drowns, in his 1983 meta-analysis of 30 studies, and Samson concluded that the greatest results from instruction in test-awareness occur with more than nine hours of total time invested.[25] Berliner, in his study of the above mentioned meta-analyses, indicated that the average student would probably score an additional 14 percentile points if given approximately 10 hours of test-wiseness training.[26] Bangert-Drowns is quick to point out that more is not necessarily better. He explains that the relation between time and test results is not a direct linear relationship, but a logarithmic one.[27] One will observe diminishing gains with increasing investments of time. Instruction time of 10 hours would be an initial investment in test-wiseness. According to Jones and Ligon, to be most effective, test-wiseness should be reviewed annually or at least just prior to each standardized testing situation.[28] The amount of time needed for the review of test-wiseness skills each year after the initial instruction needs further study.

Test-wiseness has also been studied in relation to age, sex, IQ, race, and socioeconomic level. Jongsma, in his review of 19 research studies, concluded that "test-wiseness can be effectively taught to students of all ages from preschool through adult. Al-

BEST COPY AVAILABLE

though the results have not always been statistically significant, the gains nearly always favor the group receiving instruction."[29] Samson, in his analysis of 24 studies, found "study variables such as grade level, gender, number of classes involved, number of teachers participating, ethnic background or socioeconomic level of students, the geographic area of the country, and the type of community (urban or suburban) yielded no significant differences."[30]

Sex differences in relation to test-wiseness have received little attention in research. Only four studies included this topic and their results were conflicting. Crehan, in his research of students in grades 5 through 11, found significant sex differences for grades 6 to 8 only, with males demonstrating higher test-wiseness than females.[31] Slakter found no evidence of sex, or sex by grade interaction effects on test-wiseness in his study of 5th through 11th grade students.[32] In his study of 131 kindergarten children Dreisbach found that the children trained in four test-taking skills performed better than the control group and the girls performed significantly better than the boys overall.[33] "Ferrell found no conclusive relationship between sex and test-wiseness in the high school population he studied."[34]

Sarnacki has concluded, from his review of a Dunn and Goldstein study, that the ability to pick up test-wiseness cues may be demonstrated at all levels of intelligence.[35] This finding is further emphasized in the research of Scruggs with LD and BD students. After receiving eight training sessions on test-taking skills, third and fourth grade LD and BD students scored significantly higher on the Word Study Skills subtest and exhibited descriptive increases with respect to other subtests of the Stanford Achievement Test.[36]

Very few projects have been conducted with regard to test-wiseness and race or socioeconomic level. In Urman's study of 208 Black, White, and Hispanic students in grades three and five, he found the "interaction between race and test-wiseness training was not significant."[37] Jongsma, in reviewing two research studies, concludes that "training in test-wiseness is no more effective for students of any particular socioeconomic level."[38]

Conclusion

One of the most important aspects of test-wiseness deals with its relationship to test validity. Most attacks on testing today question the validity of the tests. Jongsma points out that "virtually all authorities in educational measurement recognize test-wiseness as a source of variance on educational tests. Most would also agree that variance due to test-wiseness is undesirable in that it reduces the validity of the test."[39] These negative effects of test-wiseness are best alleviated through test-wiseness training. Training represents an attempt to ensure that all students possess a relatively equal level of the test-wiseness trait, so that the presence of test-wiseness cues on a test will not differentially reward high test-wise students while at the same time penalizing individuals low in test-wiseness. "It is also thought that the elimination or reduction in differences in test-taking skills will provide truer estimates of the knowledge and abilities being assessed,"[40] thus adding to the validity of the test. "Ebel expresses concern over excessive test-taking skills which enable some examinees to score well on a test for which they are totally unprepared. However, he considers that more error in measurement is likely to result from students who have too little rather than too much skill in taking tests."[41] Research supports some of these contentions. According to Urman, research has shown that "a lack of test-wiseness can penalize certain students and the bias against students who are not test-wise extends to standardized tests."[42] Of special importance is that if differences due to a lack of test-wiseness could be reduced, then error variance attributed to test-wiseness could be minimized and supposedly test scores would increase. Thus, as test-wiseness techniques are taught to students, a more accurate reflection or estimation of students' actual knowledge will increase and be improved.

– JoAnn M. Prell and Paul A. Prell

---

1. Pamela George, "Coaching for Tests: A Critical Look at the Issues," *Curriculum Review*, September/October 1985, p. 23.
2. Charles S. Gifford and John L. Fluitt, "Test-Taking Skills and Test Wiseness," *The Clearing House*, October 1978, p. 53.
3. Jerry A. Summers and Robert E. Shobe, *Improving Test-Taking Skills* (Terre Haute, Ind.: Curriculum Research and Development Center, Indiana State University, 1983), p. 1.
4. Gifford, p. 53.
5. Gregg W. Downey, "Is It Time We Started Teaching Children How to Take Tests?" *The American School Board Journal*, January 1977, p. 27.
6. George, p. 23.
7. *Ibid.*
8. David Savage, "Scrutinize Students' Test Scores and They Might Not Look So Rosy," *The American School Board Journal*, August 1984, p. 22.
9. George, p. 23.
10. Downey, p. 27.
11. *Ibid.*, p. 30.
12. Randolph E. Sarnacki, "An Examination of Test-Wiseness in the Cognitive Test Domain," *Review of Educational Research*, Spring 1979, p. 252.
13. Jason Millman et al., "An Analysis of Test-Wiseness," *Educational and Psychological Measurement*, Autumn 1965, p. 710.
14. Sarnacki, p. 253.
15. Millman, p. 712.
16. Valerie Ford, *Everything You Wanted to Know About Test-Wiseness* (Princeton, N.J.: Educational Testing Service, 1973), p. 10.
17. Sarnacki, p. 263.
18. *Ibid.*, p. 265.
19. Eugene Jongsma and Elaine Warshauer, *The Effects of Instruction in Test-Taking Skills Upon Student Performance on Standardized Achievement Tests* (New Orleans, La.: Louisiana Department of Elementary and Secondary Education, 1975), p. 20.
20. Phil Jones and Glynn Ligon, *Preparing Students for Standardized Tests: A Literature Review* (Austin, Tex.: Texas Office of Research and Evaluation, 1981), p. 2.
21. Sarnacki, p. 265.
22. *Ibid.*, p. 275.
23. Jones, p. 3.
24. Gordon E. Samson, "Effects of Training in Test-Taking Skills on Achievement Test Performance: A Quantitative Synthesis," *The Journal of Educational Research*, May/June 1985, p. 263.
25. Robert L. Bangert-Drowns, James A. Kulik, and Chen-Lin C. Kulik, "Effects of Coaching Programs on Achievement Test Performance," *Review of Educational Research*, Winter 1983, p. 579.
26. David Berliner, "Should Students Be Made Test Wise?" *Instructor*, February 1986, p. 23.
27. Bangert-Drowns, p. 582.
28. Jones, p. 3.

29. Jongsma, p. 18.
30. Samson, p. 262.
31. Kevin D. Crehan et al., "Longitudinal Studies of Test-Wiseness," *Journal of Educational Measurement*, Fall 1974, p. 211.
32. Malcolm J. Slakter et al., "Grade Level, Sex, and Selected Aspects of Test-Wiseness," *Journal of Educational Measurement*, Summer 1970, pp. 119-122.
33. Harold Urman, *The Effect of Test-Wiseness Training on the Achievement of Third and Fifth Grade Students* (Montreal, Quebec: The Annual Meeting of the National Council on Measurement in Education, 1983), p. 7.
34. Jongsma, p. 12.
35. Sarnacki, p. 269.
36. Thomas E. Scruggs, *The Administration and Interpretation of Standardized Achievement Tests with Learning Disabled and Behaviorally Disordered Elementary School Children* (Salt Lake City, Utah: Utah University, 1984), p. 63.
37. Urman, p. 14.
38. Jongsma, p. 13.
39. *Ibid.*, p. 2.
40. *Ibid.*, p. 6.
41. *Ibid.*, p. 5.
42. Urman, p. 2.

Bibliography

Cohen, Monroe. *Testing and Evaluation: New Views*. Washington, D.C.: Association for Childhood Education International, 1975.

Dolly, John and Kathy S. Williams. *Teaching Test-Wiseness*. Jackson Hole, Wyo.: Northern Rocky Mountain Educational Research Association, 1983.

Eakins, Darwin J., et al. "The Effects of an Instructional Test-Taking Unit on Achievement Test Scores," *The Journal of Educational Research*, November/December 1976, pp. 67-71.

Gross, Leon J. "The Effects of Three Selected Aspects of Test-Wiseness on the Standardized Test Performance of Eighth Grade Students," *Dissertation Abstracts International*, 1976, p. 6551A.

Kesselman-Turkel, Judi and Franklynn Peterson. *Test-Taking Strategies*. Chicago, Ill.: Contemporary Books, Inc., 1981.

Kulik, James, et al. "Effects of Practice on Aptitude and Achievement Test Scores," *American Educational Research Journal*, Summer 1984, pp. 435-447.

Ligon, Glynn and Phil Jones. *Preparing Students for Standardized Testing: One District's Perspective*. Austin, Tex.: Austin Independent School District, 1981.

Lyman, Howard B. *Test Scores and What They Mean*. Englewood Cliffs, N.J.: Prentice Hall, 1963.

Rawl, Edgar H. "Test-Taking Strategies Can Be the Key to Improving Test Scores," *NASSP Bulletin*, January 1984, pp. 108-112.

Eighth Street & Union Avenue, Box 789, Bloomington, Indiana 47402 Telephone 812/339-1156